

CHARACTERIZING POLYTOPES IN THE 0/1-CUBE WITH BOUNDED CHVÁTAL-GOMORY RANK

YOHANN BENCHETRIT, SAMUEL FIORINI, TONY HUYNH, AND STEFAN WELTGE

ABSTRACT. Let $S \subseteq \{0, 1\}^n$ and R be any polytope contained in $[0, 1]^n$ with $R \cap \{0, 1\}^n = S$. We prove that R has bounded Chvátal-Gomory rank (CG-rank) provided that S has bounded *pitch* and bounded *gap*, where the pitch is the minimum integer p such that all p -dimensional faces of the 0/1-cube have a nonempty intersection with S , and the gap is a measure of the size of the facet coefficients of $\text{conv}(S)$.

Let $H[\bar{S}]$ denote the subgraph of the n -cube induced by the vertices not in S . We prove that if $H[\bar{S}]$ does not contain a subdivision of a large complete graph, then both the pitch and the gap are bounded. By our main result, this implies that the CG-rank of R is bounded as a function of the treewidth of $H[\bar{S}]$. We also prove that if S has pitch 3, then the CG-rank of R is always bounded. Both results generalize a recent theorem of Cornuéjols and Lee [8], who proved that the CG-rank is always bounded if the treewidth of $H[\bar{S}]$ is at most 2.

Finally, we complement these results by proving that 0/1-polytopes $P = \text{conv}(S)$ in \mathbb{R}^n admit extended formulations whose size is bounded in terms of the pitch and the depth D of any circuit deciding membership in S . Our bound is polynomial in n whenever the pitch is constant and D is logarithmic in n .

1. INTRODUCTION

Given a polytope $R \subseteq \mathbb{R}^n$, its first *Chvátal-Gomory-closure* (*CG-closure*) is defined as $R' := \{x \in \mathbb{R}^n : c^\top x \geq \lceil \min_{y \in R} c^\top y \rceil \ \forall c \in \mathbb{Z}^n\}$, which can be shown to be again a (rational) polytope [10] with $R' \cap \mathbb{Z}^n = R \cap \mathbb{Z}^n$. By setting $R^{(0)} := R$ and $R^{(t)} := (R^{(t-1)})'$ for every $t \in \mathbb{Z}_{\geq 1}$, one recursively defines the t -th CG-closure $R^{(t)}$ of R . It is well-known that there exists a number $t \in \mathbb{Z}_{\geq 0}$ such that $R^{(t)} = \text{conv}(R \cap \mathbb{Z}^n)$, and the smallest such number is called the *Chvátal-Gomory-rank* (*CG-rank*) of R . In this paper, we give new bounds on the CG-rank of a polytope R contained in $[0, 1]^n$ that only depend on properties of $S = R \cap \{0, 1\}^n$ and not on R itself. This is in the spirit of [7] except that we only consider relaxations contained in $[0, 1]^n$.

One particular reason to study the CG-rank is to obtain bounds on lengths of *cutting-plane proofs* as introduced in [4, Sec. 6]. Letting k be the CG-rank of $R \subseteq \mathbb{R}^n$, the length of a cutting-plane proof is at most $(n^{k+1} - 1)/(n - 1)$. In fact, if k is a fixed constant and $R \subseteq \mathbb{R}^n$ has CG-rank k , then optimizing a linear function over $R \cap \mathbb{Z}^n$ is one of the few problems that is known to be in $\text{coNP} \cap \text{NP}$ but not known to be in P , see for instance [9]. While the CG-rank of general polytopes in \mathbb{R}^n can be arbitrarily large compared to n (even for $n = 2$), the CG-rank of a polytope contained in $[0, 1]^n$ is always bounded by $\mathcal{O}(n^2 \log n)$, see [12]. Unfortunately, there exist polytopes in $[0, 1]^n$ whose CG-rank grows quadratically in n , see [17].

This motivates the study for situations in which the CG-rank is at most a constant independent of n . This question has been recently addressed by Cornuéjols & Lee [8]. In their work, given a set $S \subseteq \{0, 1\}^n$, they consider the graph $H[\bar{S}]$ whose vertices are the points of $\bar{S} := \{0, 1\}^n \setminus S$ where two points are adjacent if they differ in exactly one coordinate. Their main result is that if the treewidth of $H[\bar{S}]$ (denoted $\text{tw}(H[\bar{S}])$) is at most 2, then the CG-rank of any polytope $R \subseteq [0, 1]^n$ with $R \cap \{0, 1\}^n = S$ is bounded (they prove a bound of 4, which is tight). One corollary of our work is that this holds for *all* values of treewidth: the CG-rank of every polytope $R \subseteq [0, 1]^n$ with $R \cap \{0, 1\}^n = S$ is bounded by a function that only depends on the treewidth of $H[\bar{S}]$.

In order to state our main result, we define the *pitch* of a subset $S \subseteq \{0, 1\}^n$ as the smallest $p \in \mathbb{Z}_{\geq 0}$ such that every p -dimensional face of $[0, 1]^n$ has a nonempty intersection with S . If S is empty, we define $p := n + 1$. We remark that this definition of pitch is consistent with the original definition due to Bienstock & Zuckerberg [3]. Furthermore, we define the *gap* of

S as the smallest $\Delta \in \mathbb{Z}_{\geq 0}$ such that $\text{conv}(S)$ can be described as the set of solutions $x \in \mathbb{R}^n$ satisfying inequalities of the form

$$(1) \quad \sum_{i \in I} c_i x_i + \sum_{j \in J} c_j (1 - x_j) \geq \delta$$

where I, J are disjoint subsets of $[n]$, $\delta, c_1, \dots, c_n \in \mathbb{Z}_{\geq 0}$ with $\delta \leq \Delta$. We require that for each inequality in the description, the corresponding equation (obtained from (1) by replacing the inequality sign by an equality sign) defines hyperplane spanned by 0/1-points. Notice that if S is empty, then we have $\Delta = 1$.

The gap is well-defined since for every $S \subseteq \{0, 1\}^n$, $\text{conv}(S)$ has a description by inequalities in which every corresponding hyperplane is generated by 0/1-points. To see this, consider a full-dimensional 0/1- polytope $\text{conv}(T)$ where $S \subseteq T \subseteq \{0, 1\}^n$ such that $\text{conv}(S)$ is a face of $\text{conv}(T)$ (this exists since the set $\{0, 1\}^n$ is full-dimensional). Clearly, the bounding hyperplane of every facet of $\text{conv}(T)$ is generated by 0/1-points. Since $\text{conv}(S)$ is the intersection of the facets of $\text{conv}(T)$ which contain it, the claimed description directly follows.

Our main result is the following.

Theorem 1. *Let $S \subsetneq \{0, 1\}^n$ be a set with pitch p and gap Δ . Then the CG-rank of any polytope $R \subseteq [0, 1]^n$ with $R \cap \{0, 1\}^n = S$ is at most $p + \Delta - 1$.*

In order to generalize the result of Cornuéjols & Lee, we will show that p and Δ are both bounded in terms of $\text{tw}(H[\bar{S}])$. In fact, we will not even need the definition of treewidth because we actually prove a stronger result. We let K_t be a clique on t vertices. A *subdivision* of K_t is a graph obtained from K_t by replacing each edge of K_t by a path.

Corollary 2. *Let $S \subseteq \{0, 1\}^n$ and let t be the maximum integer such that $H[\bar{S}]$ contains a subdivision of K_{t+1} . Then the CG-rank of any polytope $R \subseteq [0, 1]^n$ with $R \cap \{0, 1\}^n = S$ is at most $t + 2t^{1/2}$.*

To see that Corollary 2 is a generalization of the result by Cornuéjols & Lee, the only thing the reader needs to know is that if a graph G has a subdivision of K_{t+1} , then $\text{tw}(G) \geq t$. This is an easy fact (see [11] for a gentle introduction to treewidth).

We also discuss further properties of sets $S \subseteq \{0, 1\}^n$ whose pitch is small. For instance, if we denote by p the pitch of S , we observe that optimizing a linear function over S can be done with $\mathcal{O}(n^p)$ oracle calls using an oracle that decides if a point $x \in \{0, 1\}^n$ belongs to S , see Proposition 4. This algorithm already appears in [8], but we show that it is valid under a weaker hypothesis. Furthermore, we show that $\text{conv}(S)$ has an extended formulation of size $\mathcal{O}(n2^{pD})$, where D is the depth of any Boolean circuit of fan-in 2 deciding S , see Section 6.

Paper structure. In Section 2 we discuss the meaning of the parameters p and Δ , and in particular their relation to the CG-rank. For instance, we give examples that show that the CG-rank of a polytope in $[0, 1]^n$ cannot be bounded in only one of the two parameters. Section 3 contains the proof of Theorem 1. In Section 4 we complement our main theorem by a result quantifying how well the t -th CG-closure approximates $\text{conv}(S)$ for constant t and constant p , this time without bounding Δ . In Section 5 we investigate the convex hulls of sets with pitch $p = 3$. In this case, we show that Δ is automatically bounded and give a complete linear description of $\text{conv}(S)$. We show that treewidth at most 2 implies pitch at most 3, but not vice versa, hence this result also strictly generalizes the main result of Cornuéjols & Lee [8]. Finally, extended formulations of $\text{conv}(S)$ for small pitch sets S are discussed in Section 6.

2. DISCUSSION OF THE PARAMETERS

In this section, we discuss how the parameters *pitch* and *gap* of a set $S \subseteq \{0, 1\}^n$ influence the CG-rank of polytopes $R \subseteq [0, 1]^n$ with $R \cap \{0, 1\}^n = S$.

2.1. Small CG-rank implies small pitch. We first observe that, in order to get a constant bound on the CG-rank, one has to restrict to sets S with bounded pitch. Although this follows directly from known results, we include a proof for completeness.

Proposition 3. *Let $S \subseteq \{0, 1\}^n$ have pitch p . Then there exists a polytope $R \subseteq [0, 1]^n$ with $R \cap \{0, 1\}^n = S$ whose CG-rank is at least $p - 1$.*

Proof. Following [8], we let R be the worst possible¹ relaxation of $\text{conv}(S)$:

$$(2) \quad R := \left\{ x \in [0, 1]^n \mid \forall a \in \bar{S} : \sum_{i:a_i=0} x_i + \sum_{i:a_i=1} (1 - x_i) \geq \frac{1}{2} \right\}.$$

By the definition of p , there exists a $(p - 1)$ -dimensional face F of $[0, 1]^n$ such that $F \cap S = \emptyset$. The CG-rank of R is at least that of its face $F \cap R$ (see for instance [6, Lem. 5.17]), which can be shown to be exactly $p - 1$ using [4, Lem. 7.2]. \square

It turns out that the structure of sets $S \subseteq \{0, 1\}^n$ with small pitch p can be efficiently exploited with respect to certain optimization tasks. For instance, the p -th level of the Bienstock-Zuckerberg hierarchy [3] gives a tight description of $\text{conv}(S)$, at least when applied to sets S of set-covering type. A much simpler observation is that linear programming over S is easy if p is constant.

2.2. Optimization algorithm for small pitch. Let $S \subseteq \{0, 1\}^n$ have pitch p . Assume that we have an oracle for deciding if a given point $x \in \{0, 1\}^n$ belongs to S . Here, we prove that optimizing a linear function over S can be done after performing at most $\mathcal{O}(n^p)$ oracle calls, and spending an extra polynomial time to select an optimum solution.

The algorithm is as follows. Given a cost vector $c \in \mathbb{R}^n$, we let $x^* \in \{0, 1\}^n$ be defined as $x_i^* := 0$ if $c_i \geq 0$ and $x_i^* := 1$ if $c_i < 0$. Note that this is an optimum solution of $\min\{c^\top x \mid x \in \{0, 1\}^n\}$. Next, among all the vertices of the cube $x \in \{0, 1\}^n$ that are at Hamming distance at most p from x^* , output any vertex x that belongs to S and has minimum cost.

Proposition 4. *For every $S \subseteq \{0, 1\}^n$ with pitch p and every $c \in \mathbb{R}^n$, the algorithm described above solves $\min\{c^\top x \mid x \in S\}$ in $\mathcal{O}(n^p)$ oracle calls.*

Proof. Clearly, the number of oracle calls performed by the algorithm is at most

$$\sum_{k=0}^p \binom{n}{k} \leq (n + 1)^p = \mathcal{O}(n^p).$$

There is always a feasible solution $x \in \{0, 1\}^n$ at Hamming distance at most p from x^* , since otherwise there would exist a p -dimensional face of the cube that is disjoint from S , which contradicts that the pitch of S is p . Therefore, the algorithm always outputs some feasible solution.

In order to finish proving the correctness of the algorithm, consider an optimum solution x^{opt} in S whose Hamming distance $d_H(x^{\text{opt}}, x^*)$ to x^* is minimum. Let $I := \{i \in [n] \mid x_i^{\text{opt}} \neq x_i^*\}$ be the set of indices of bits of x^* that are flipped in x^{opt} , so that we can express the optimum value as

$$\text{OPT} = c^\top x^{\text{opt}} = c^\top x^* + \sum_{i \in I} |c_i|.$$

Now consider the set F of vertices $x \in \{0, 1\}^n$ that are obtained by flipping the bits of x^* indexed by some set $J \subseteq I$. Thus, $F = \{x \in \{0, 1\}^n \mid \forall i \in [n] \setminus I : x_i = x_i^* = x_i^{\text{opt}}\}$. Clearly, F is the vertex set of some face of the cube. Every $x \in F$ has cost at most OPT since we have

$$c^\top x = c^\top x^* + \sum_{i \in J} |c_i| \leq c^\top x^* + \sum_{i \in I} |c_i| = c^\top x^{\text{opt}} = \text{OPT}.$$

By minimality of $d_H(x^{\text{opt}}, x^*)$, no $x \in F \setminus \{x^{\text{opt}}\}$ belongs to S . Thus, $d_H(x^{\text{opt}}, x^*) \leq p$ (otherwise, F would contain a p -dimensional face of $[0, 1]^n$ disjoint from S), and x^{opt} is one of the feasible solutions considered by the algorithm. The result follows. \square

¹In the sense that $R' \supseteq Q'$ for all polytopes $Q \subseteq [0, 1]^n$ such that $Q \cap \{0, 1\}^n = R \cap \{0, 1\}^n = S$.

2.3. Small CG-rank implies small gap. One might wonder whether sets $S \subseteq \{0, 1\}^n$ with small pitch are already simple enough to ensure that every relaxation for S contained in $[0, 1]^n$ has small CG-rank. However, it turns out that such sets S also need to have a description with bounded coefficients only, as illustrated by the next two results. Here, we denote by $\|A\|_\infty$ the maximum absolute value of an entry of A .

Lemma 5. *Let $P = \{x \in \mathbb{R}^n \mid Ax \geq b\}$, where $A \in \mathbb{Z}^{m \times n}$ and $b \in \mathbb{Z}^m$. Letting P' denote the first CG-closure of P , there is a description $P' = \{x \in \mathbb{R}^n \mid Bx \geq c\}$ with B and c integer such that $\|B\|_\infty \leq n\|A\|_\infty$.*

Proof. Every valid inequality for P' can be written as $\lambda^\top Ax \geq \lceil \lambda^\top b \rceil$ for some $\lambda \in \mathbb{R}_+^m$. By Carathéodory's theorem, we may assume that λ has at most n non-zero entries. Furthermore, it is well known that one can replace every entry of λ by its non-integral part to obtain an inequality that is valid for P' and at least as strong as the original one (see, e.g., [6, Lem. 5.13]). In other words, we may assume that $\lambda \in [0, 1]^m$ and λ has at most n non-zero entries. By the triangle inequality, this implies

$$\|\lambda^\top A\|_\infty = \left\| \sum_{i: \lambda_i \neq 0} \lambda_i A_i \right\|_\infty \leq \sum_{i: \lambda_i \neq 0} \underbrace{\lambda_i \|A_i\|_\infty}_{\leq \|A\|_\infty} \leq n\|A\|_\infty,$$

and the lemma follows. \square

Proposition 6. *Let $S \subsetneq \{0, 1\}^n$ be nonempty with gap Δ . Then there exists a polytope $R \subseteq [0, 1]^n$ with $R \cap \{0, 1\}^n = S$ such that the CG-rank of R is at least $\frac{\log \Delta}{\log n} - 1$.*

Proof. First, we claim that every integer matrix A for which there is an integer vector b with $\text{conv}(S) = \{x \in \mathbb{R}^n \mid Ax \geq b\}$ satisfies $\|A\|_\infty \geq \frac{\Delta}{n}$. Indeed, every inequality in such a description is of the form

$$\sum_{i \in I} c_i x_i - \sum_{j \in [n] \setminus I} c_j x_j \geq \beta$$

where $I \subseteq [n]$, $c \in \mathbb{Z}_{\geq 0}^n$, $\beta \in \mathbb{Z}$. Letting $\delta := \beta + \sum_{j \in [n] \setminus I} c_j$ we can rewrite this inequality as

$$\sum_{i \in I} c_i x_i + \sum_{j \in [n] \setminus I} c_j (1 - x_j) \geq \delta.$$

By the definition of Δ , for at least one of these inequalities we must have $\delta = \Delta$. Since $\text{conv}(S) \subseteq [0, 1]^n$ is nonempty, this implies $\|c\|_\infty \geq \frac{\Delta}{n}$.

Second, consider the polytope $R := \{x \in [0, 1]^n \mid \forall a \in \bar{S} : \sum_{i: a_i=0} x_i + \sum_{j: a_j=1} (1 - x_j) \geq 1\}$. Note that $R \subseteq [0, 1]^n$ and $R \cap \{0, 1\}^n = S$. Furthermore, R has a description of the form $R = \{x \in \mathbb{R}^n \mid Cx \geq d\}$ with C, d integer and $\|C\|_\infty = 1$. Thus, letting k be the CG-rank of R and in view of Lemma 5, we obtain $n^k \geq \frac{\Delta}{n}$, which yields the claim. \square

Let $S \subseteq \{0, 1\}^n$ with pitch p and gap Δ , and denote by k the largest CG-rank of a polytope $R \subseteq [0, 1]^n$ with $R \cap S = \{0, 1\}^n$. Summarizing the previous observations, we have seen that k can be bounded from below in terms of p (Proposition 3), and also in terms of Δ and n (Proposition 6). This explains the occurrence of both parameters in the statement of Theorem 1.

In what follows next, we would like to discuss that none of the two parameters p and Δ can be bounded by a function that only depends on the other. To see that p cannot be bounded by a function in Δ , observe that the set $S = \{x \in \{0, 1\}^n \mid x_p + x_{p+1} + \dots + x_n \geq 1\}$ has pitch p and gap 1.

2.4. Bounded Pitch Does Not Imply Bounded CG-rank. Next, we show that neither the parameter Δ nor the CG-rank can be bounded in terms of p alone.

Proposition 7. *For each $n \in \mathbb{N}$, there exists $S_n \subseteq \{0, 1\}^{2^{n+2}}$ such that S_n has pitch at most 7 but gap at least 2^{n+1} .*

Proof. Fix $n \in \mathbb{N}$. We define a vector $c \in \mathbb{R}^{2n+2}$ by setting $c_1 = 2^n$, $c_2 = 2^{n-1}$, $c_i = c_{i-1}$ if $i \in [3, 2n+1]$ is odd, $c_i = (2^n - c_{i-1})/2$ if $i \in [3, 2n+1]$ is even, and $c_{2n+2} = 2^n - c_{2n+1}$. Now consider the inequality $\sum_{i=1}^{2n+2} c_i x_i \geq 2^{n+1}$, and let S_n be the set of vectors in $\{0, 1\}^{2n+2}$ for which this inequality is satisfied.

By definition, $\sum_{i=1}^{2n+2} c_i x_i \geq 2^{n+1}$ is a valid inequality for $\text{conv}(S_n)$. We claim that it is actually a facet of $\text{conv}(S_n)$. This follows by observing that $c_1 + c_2 + c_3 = 2^{n+1}$, $c_1 + c_{2i-1} + c_{2i} + c_{2i+1} = c_1 + c_{2i-2} + c_{2i} + c_{2i+1} = 2^{n+1}$ for all $i \in [2, n]$, $c_1 + c_{2n} + c_{2n+2} = c_1 + c_{2n+1} + c_{2n+2} = 2^{n+1}$ and $c_2 + c_3 + c_{2n+1} + c_{2n+2} = 2^{n+1}$.

Note that the greatest common divisor of the entries of c is 1 since c_1 is a power of 2 and c_{2n} is odd. Since all c_i are non-negative, this implies that the gap of S_n is at least 2^{n+1} .

Finally, we show that S_n has pitch at most 7. That is, we must show that the 7 smallest entries of c sum to at least 2^{n+1} . This is easily checked by hand if $n < 8$, so we may assume $n \geq 8$. By solving a linear recurrence of degree 1, we find that $c_{2i} = c_{2i+1} = 2^n \cdot (1 - (-1/2)^i)/3$ for $i \in [1, n]$. It follows that the 7 smallest entries of c are $c_4, c_5, c_8, c_9, c_{12}, c_{13}$, and c_{16} . The sum of these entries is

$$2^n \cdot \left(\frac{1}{4} + \frac{1}{4} + \frac{5}{16} + \frac{5}{16} + \frac{21}{64} + \frac{21}{64} + \frac{85}{256} \right) = 2^n \cdot \frac{541}{256} > 2^n \cdot 2 = 2^{n+1},$$

as required. \square

By Proposition 6 and Proposition 7 we directly obtain.

Corollary 8. *For each n , there exists a polytope $R \subseteq [0, 1]^{2n+2}$ such that $S = R \cap \{0, 1\}^{2n+2}$ has pitch at most 7, but the CG-rank of R is $\Omega(\frac{n}{\log n})$.*

2.5. Bounded Treewidth Implies Bounded Pitch and Gap. Finally, we demonstrate that Theorem 1 can indeed be seen as a generalization of the results of Cornuéjols & Lee [8]. To this end, it suffices to show that p and Δ can be bounded in terms of the treewidth of $H[\bar{S}]$. Recall that the largest t such that $H[\bar{S}]$ contains a subdivision of K_{t+1} is at most $\text{tw}(H[\bar{S}])$.

Lemma 9. *Let $S \subseteq \{0, 1\}^n$, and let p and Δ respectively denote the pitch and the gap of S . If t is maximum such that $H[\bar{S}]$ contains a subdivision of K_{t+1} , then $p \leq t + 1$ and $\Delta \leq 2t^{t/2}$.*

Proof. Note that the d -dimensional cube contains a subdivision of K_{d+1} , where the branch vertices are the vectors with support at most 1, and the subdivision vertices are the vectors with support 2. Now, since $H[\bar{S}]$ contains a subgraph isomorphic to the $(p-1)$ -dimensional cube, it contains a subdivision of K_p and we have $t \geq p-1$.

To show $\Delta \leq 2t^{t/2}$, observe that it suffices to prove the following. For any hyperplane $H := \{x \in \mathbb{R}^n \mid \sum_{i \in I} c_i x_i + \sum_{j \in [n] \setminus I} c_j (1 - x_j) = 1\}$ that is spanned by 0/1-points such that $\sum_{i \in I} c_i x_i + \sum_{j \in [n] \setminus I} c_j (1 - x_j) \geq 1$ is valid for S and $c_1, \dots, c_n \in \mathbb{Q}_{\geq 0}$, there exists some integer number $K \in [1, 2t^{t/2}]$ such that every c_i is an integer multiple of $1/K$.

By switching the coordinates indexed by $[n] \setminus I$, we may assume that $I = [n]$. Define $I_{<1/2} := \{i \in [n] \mid c_i < 1/2\}$, and $I_{=1/2}, I_{>1/2}$ similarly. We have that $|I_{<1/2}| \leq t$ since otherwise $H[\bar{S}]$ contains a subdivision of a clique of size $t+2$ whose branch vertices are the characteristic vectors of the empty set \emptyset and the singletons $\{i\}$ for $i \in I_{<1/2}$ and whose subdivision vertices are the characteristic vectors of the pairs $\{i, j\}$ for $i, j \in I_{<1/2}$.

Let $x \in \{0, 1\}^n \cap H$ and denote by T its support. Then one of the following holds: (i) $|T \cap I_{=1/2}| = |T \cap I_{>1/2}| = 0$, (ii) $|T \cap I_{=1/2}| = 1$ and $|T \cap I_{>1/2}| = 0$, (iii) $|T \cap I_{=1/2}| = 0$ and $|T \cap I_{>1/2}| = 1$, or (iv) $|T \cap I_{=1/2}| = 2$ and $|T \cap I_{>1/2}| = 0$. Thus, the vector c is the unique solution of a system of linear equations of the following form

$$\begin{pmatrix} A & & \\ B & * & \\ C & & D \\ & \mathbb{I} & \end{pmatrix} c = b,$$

where the coefficient matrix has integer entries, A, B, C are 0/1-matrices with columns indexed by $I_{<1/2}$, \mathbb{I} is an identity matrix with columns indexed by $I_{=1/2}$, D is a 0/1-matrix with columns

indexed by $I_{>1/2}$ and exactly one 1 per row, and b is a column vector with entries in $\{1/2, 1\}$. The last rows of the above system are meant to be the trivial equations $c_i = 1/2$, which are obviously valid for all $i \in I_{=1/2}$. Since every row in D contains exactly one 1, we can perform elementary row operations to obtain an equivalent system of the form

$$\begin{pmatrix} E & * \\ * & \mathbb{I} \\ & & \mathbb{I} \end{pmatrix} c = b',$$

where the coefficient matrix has integer entries, E is a matrix with entries in $\{-1, 0, 1\}$ and columns indexed by $I_{<1/2}$, and b' a column vector with entries in $\{0, 1/2, 1\}$. By removing some rows in the topmost block, we may assume that the coefficient matrix is a regular $n \times n$ -matrix whose determinant is $\pm \det(E)$. Thus, by Cramer's rule, every c_i is an integer multiple of $\frac{1}{2|\det(E)|}$. Since E is a matrix with entries in $\{-1, 0, 1\}$ and $|I_{<1/2}| \leq t$ columns and rows, by the Hadamard bound we obtain

$$K = 2|\det(E)| \leq 2t^{\frac{1}{2}t},$$

as claimed. \square

3. PROOF OF MAIN THEOREM

Lemma 10. *Let $R \subseteq [0, 1]^n$ be a polytope and $I, J \subseteq [n]$ with $I \cap J = \emptyset$ such that*

$$(3) \quad \sum_{i \in I} x_i + \sum_{j \in J} (1 - x_j) \geq 1$$

holds for every $x \in R \cap \{0, 1\}^n$. Then (3) is also valid for $R^{(n+1-(|I|+|J|))}$.

Proof. Consider the set

$$F := \{x \in R \mid x_i = 0 \ \forall i \in I, x_j = 1 \ \forall j \in J\},$$

which is a face of R of dimension $k \leq n - (|I| + |J|)$. Since (3) is valid for $R \cap \{0, 1\}^n$, we have $F \cap \mathbb{Z}^n = \emptyset$. Since $F \subseteq [0, 1]^n$, this implies $F^{(k)} = \emptyset$ (by [12, Lem. 2.2]). This implies $R^{(k)} \cap F = \emptyset$ (see [6, Lem. 5.17]) and hence there exists an $\varepsilon > 0$ such that

$$\sum_{i \in I} x_i + \sum_{j \in J} (1 - x_j) \geq \varepsilon$$

is valid for $R^{(k)}$. This means that (3) holds for $R^{(k+1)}$, as claimed. \square

Proof of Theorem 1. By the definition of Δ , we can find a description of $\text{conv}(R \cap \{0, 1\}^n)$ by means of linear inequalities where every inequality is of the form

$$(4) \quad \sum_{i \in I} c_i x_i + \sum_{j \in J} c_j (1 - x_j) \geq \delta$$

for some $I, J \subseteq [n]$ with $I \cap J \neq \emptyset$, where $\delta \in \mathbb{Z}_{\geq 0}$, $c_i \in \mathbb{Z}_{\geq 1}$ for all $i \in I \cup J$, and $\delta \leq \Delta$. Note that every such inequality with $\delta = 0$ is already valid for R . For inequalities with $\delta \geq 1$, we may assume that $c_i \leq \delta$ holds for every $i \in I \cup J$.

By induction on $\delta \geq 1$ we will show that (4) holds for every $x \in R^{(p+\delta-1)}$, which then yields the claim. If $\delta = 1$, then we have $c_i = 1$ for all $i \in I \cup J$. By Lemma 10, we know that Inequality (4) is valid for $R^{(t)}$, where $t = n + 1 - (|I| + |J|)$. It remains to show that $t \leq p + \delta - 1 = p$. To this end, consider the set

$$F = \{x \in [0, 1]^n \mid x_i = 0 \ \forall i \in I, x_j = 1 \ \forall j \in J\},$$

which is a face of the cube, and note that no point of F satisfies (4). Thus, we indeed obtain $p \geq \dim(F) + 1 = n + 1 - (|I| + |J|) = t$.

Now let $\delta \geq 2$. We may assume that $|I| + |J| \geq 1$, otherwise we can divide (4) by δ and proceed by induction. For every $i_0 \in I$ consider the inequality

$$\sum_{i \in I \setminus \{i_0\}} c_i x_i + (c_{i_0} - 1)x_{i_0} + \sum_{j \in J} c_j (1 - x_j) \geq \delta - x_{i_0} \geq \delta - 1,$$

which is valid for $R \cap \{0, 1\}^n$. Similarly, for every $j_0 \in J$ the inequality

$$\sum_{i \in I} c_i x_i + \sum_{j \in J \setminus \{j_0\}} c_j (1 - x_j) + (c_{j_0} - 1)(1 - x_{j_0}) \geq \delta - (1 - x_{j_0}) \geq \delta - 1$$

is also valid for $R \cap \{0, 1\}^n$. Thus, by the induction hypothesis, both such inequalities are valid for $R^{(p+\delta-2)}$. Summing these $k := |I| + |J|$ many inequalities up and dividing them by $k \geq 1$, we obtain that

$$\sum_{i \in I} \left(c_i - \frac{1}{k}\right) x_i + \sum_{j \in J} \left(c_j - \frac{1}{k}\right) (1 - x_j) \geq \delta - 1$$

is valid for $R^{(p+\delta-2)}$. Choose $\varepsilon > 0$ such that $(1 + \varepsilon)(c_i - \frac{1}{k}) \leq c_i$ holds for all $i \in I \cup J$. Scaling the above inequality by $(1 + \varepsilon)$, we thus obtain that

$$\begin{aligned} \sum_{i \in I} c_i x_i + \sum_{j \in J} c_j (1 - x_j) &\geq (1 + \varepsilon) \left(\sum_{i \in I} \left(c_i - \frac{1}{k}\right) x_i + \sum_{j \in J} \left(c_j - \frac{1}{k}\right) (1 - x_j) \right) \\ &\geq (1 + \varepsilon)(\delta - 1), \end{aligned}$$

holds for every $x \in R^{(p+\delta-2)}$, and hence

$$\sum_{i \in I} c_i x_i + \sum_{j \in J} c_j (1 - x_j) \geq \lceil (1 + \varepsilon)(\delta - 1) \rceil \geq \delta$$

is valid for $R^{(p+\delta-1)}$, as claimed. \square

4. APPROXIMATING THE INTEGER HULL WHEN THE PITCH IS BOUNDED

We have shown in Section 2.4 that if we only assume that p is constant, it might take $\Omega(n/\log n)$ rounds of CG-cuts to converge to the integer hull: we have to control Δ also in order to guarantee bounded CG-rank. Here we prove that bounding p alone is in fact enough to obtain good *approximations* of the integer hull after a bounded number of rounds. This is in contrast with the results of Singh & Talwar [18], who show that for many problems performing a constant number of rounds of CG-cuts does not significantly decrease the integrality gap.

Corollary 11. *Let $S \subseteq \{0, 1\}^n$ have pitch p and let $\varepsilon \in (0, 1)$ be such that $p\varepsilon^{-1} \in \mathbb{Z}_{\geq 0}$. For every $t \geq p\varepsilon^{-1} - 1$ and for every inequality $\sum_{i \in I} c_i x_i + \sum_{j \in J} c_j (1 - x_j) \geq \delta$ that is valid for $\text{conv}(S)$ with $\delta \geq c_1, \dots, c_n \geq 0$, the inequality $\sum_{i \in I} c_i x_i + \sum_{j \in J} c_j (1 - x_j) \geq (1 - \varepsilon)\delta$ is valid for $R^{(t)}$, where $R \subseteq [0, 1]^n$ is any polytope such that $R \cap \{0, 1\}^n = S$.*

Proof. After flipping some coordinates, we may assume that $J = \emptyset$. After scaling, we may further assume that $\delta = 1$. Let $K := p\varepsilon^{-1}$. Consider the valid inequality $\sum_{i \in I} \tilde{c}_i x_i \geq \tilde{\delta}$ where $\tilde{c}_i := \frac{1}{K} \lfloor K c_i \rfloor \in \{0, 1/K, 2/K, \dots, 1\}$ and $\tilde{\delta} := \min\{\sum_{i \in I} \tilde{c}_i x_i \mid x \in S\}$. We claim that $\tilde{\delta} \leq 1 - \varepsilon$. Indeed, let $x \in S$ be arbitrary and let $y \in S$ be such that $0 \leq y \leq x$ and y has support on at most p coordinates. Then

$$\begin{aligned} \sum_{i \in I} \tilde{c}_i x_i &\geq \sum_{i \in I} \tilde{c}_i y_i \\ &= \sum_{i \in I} \frac{1}{K} \lfloor K c_i \rfloor y_i \\ &\geq \sum_{i \in I} \frac{1}{K} (K c_i - 1) y_i \\ &\geq \sum_{i \in I} c_i y_i - \frac{p}{K} \\ &\geq 1 - \frac{p}{K} = 1 - \varepsilon \end{aligned}$$

so that $\delta \leq 1 - \varepsilon$. Now consider the valid inequality $\sum_{i \in I} K \tilde{c}_i x_i \geq K(1 - \varepsilon) = K - p$ with nonnegative integer coefficients. From the proof of Theorem 1, we see that this inequality is valid for the t -th CG-closure of R since $t = K - 1 = (K - p) + p - 1$. \square

5. THE PITCH-3 CASE

Theorem 12. *Let $S \subseteq \{0, 1\}^n$ have pitch $p \leq 3$. Then $P = \text{conv}(S)$ can be defined by $0 \leq x_i \leq 1$ for $i \in [n]$ together with inequalities that can be brought in the following form after flipping some coordinates, where for each inequality the subsets of indices are a partition of $[n]$ (we allow empty sets in the partition):*

$$\begin{aligned}
 (5) \quad & \sum_{i \in I_0} 0x_i + \sum_{i \in I_1} 1x_i \geq 1, & |I_0| \leq 2 \\
 (6) \quad & \sum_{i \in I_0} 0x_i + \sum_{i \in I_1} 1x_i + \sum_{i \in I_2} 2x_i \geq 2, & |I_0| \leq 1 \\
 (7) \quad & \sum_{i \in I_1} 1x_i + \sum_{i \in I_2} 2x_i + \sum_{i \in I_3} 3x_i \geq 3, & |I_1| \geq 3 \\
 (8) \quad & \sum_{i \in I_1} 1x_i + \sum_{i \in I_2} 2x_i + \sum_{i \in I_3} 3x_i + \sum_{i \in I_4} 4x_i \geq 4, & |I_1| = 2, |I_2| \geq 1 \\
 (9) \quad & \sum_{i \in I_2} 2x_i + \sum_{i \in I_3} 3x_i + \sum_{i \in I_4} 4x_i + \sum_{i \in I_6} 6x_i \geq 6, & |I_2| \geq 3
 \end{aligned}$$

In particular, S has gap $\Delta \leq 6$.

Proof. We may assume that $n \geq 3$, otherwise the theorem holds trivially. Thus, S is nonempty. Pick any nonredundant inequality description of $\text{conv}(S)$ such that the corresponding hyperplanes are spanned by 0/1-points. Let $c^\top x \geq \delta$ be any inequality in this description which is not of the form $x_i \geq 0$ or $1 - x_i \geq 0$. By flipping coordinates and scaling we may assume that $c_i \in \mathbb{Q}_{\geq 0}$ and $\delta = 1$. We choose a non-redundant system that uniquely defines c consisting of equations of the form $c_i^* = 0$, $c_i^* - c_j^* = 0$, and $\sum_{i \in I \subseteq [n]} c_i^* = 1$ such that equations of lower support are always included before equations of higher support. In particular, this implies that equations of the form $c_i^* = 0$ or $c_i^* = 1$ are always included if $c_i = 0$ or $c_i = 1$.

Sort the entries of c as $c_1 \leq c_2 \leq \dots \leq c_n$. Clearly, since S has pitch at most 3, $c_1 + c_2 + c_3 \geq 1$. Hence, if any equation has support greater than 3, then it is already implied by $c_1^* + c_2^* + c_3^* = 1$ together with equations of the form $c_i^* = 0$. Thus, no equation with support greater than 3 appears. If $c_1 + c_2 + c_3 = 1$, then any equation whose support has size 3 is already implied by the equation $c_1^* + c_2^* + c_3^* = 1$ together with equations of the form $c_\ell^* - c_m^* = 0$ for $\ell \in [3]$. If $c_1 + c_2 + c_3 > 1$, then no equation of the form $c_i^* + c_j^* + c_k^* = 1$ will appear. Thus, at most one equation of support 3 appears.

Define a graph $G = ([n], E)$, where $E := \{ij : c_i^* + c_j^* = 1 \text{ or } c_i^* - c_j^* = 0\}$. Let $\Sigma := \{ij : c_i^* + c_j^* = 1\}$ and define a cycle of G to be *unbalanced* if it contains an odd number of edges of Σ . Since the system is non-redundant, each component of G contains at most one cycle, which will have to be unbalanced. For each $\gamma \in [0, 1]$, let $J_\gamma := \{i \in [n] : c_i = \gamma\}$. Note that $|J_0| \leq 2$, as $c_1 + c_2 + c_3 \geq 1$. Let $J'_{\frac{1}{2}}$ be the set of vertices of G contained in a component with an unbalanced cycle. Clearly, $J'_{\frac{1}{2}} \subseteq J_{\frac{1}{2}}$.

Let T_1, \dots, T_ℓ be the components of G which contain at least one edge and no cycles. Note that if $\ell \geq 2$, then the set of solutions of the system (minus the single equation of the form $c_i^* + c_j^* + c_k^* = 1$) has dimension at least 2. Thus, the solution set of the full system has dimension at least 1, which contradicts the uniqueness of c . Therefore, $\ell \leq 1$. We may partition the vertices of T_1 as $J'_\alpha \cup J'_\beta$ where $c_i = \alpha$ for all $i \in J'_\alpha$ and $c_i = 1 - \alpha := \beta$ for all $i \in J'_\beta$. Note that if $\alpha = 0$, then $J'_\alpha \subseteq J_0$ and $J'_\beta \subseteq J_1$, and if $\alpha = \frac{1}{2}$, then $J'_\alpha \cup J'_\beta \subseteq J_{\frac{1}{2}}$.

It follows that $[n] := J_0 \cup J_\alpha \cup J_{\frac{1}{2}} \cup J_\beta \cup J_1$, for some $0 < \alpha < \frac{1}{2}$ and $\beta := 1 - \alpha$ (some of these sets are possibly empty). There are now various cases to consider depending on where the indices of the single equation $c_i^* + c_j^* + c_k^* = 1$ belong.

First suppose that there does not exist an equation of the form $c_i^* + c_j^* + c_k^* = 1$. In this case, by the uniqueness of c , we must have $J_\alpha = J_\beta = \emptyset$. If $|J_0| = 2$, then $J_{\frac{1}{2}} = \emptyset$ and $J_1 \neq \emptyset$, so we get (5) with $(I_0, I_1) = (J_0, J_1)$. If $|J_0| \leq 1$, we get (6) with $(I_0, I_1, I_2) = (J_0, J_{\frac{1}{2}}, J_1)$.

We may hence assume there does exist an equation of the form $c_i^* + c_j^* + c_k^* = 1$ (with $i < j < k$). We may further assume that $\{i, j, k\} \cap J_0 = \emptyset$, because otherwise, the equation $c_i^* + c_j^* + c_k^* = 1$ is implied by the lower support equations $c_i^* = 0$ and $c_j^* + c_k^* = 1$. Similarly, $\{i, j, k\} \cap J_1 = \emptyset$.

Suppose $\{i, j, k\} \subseteq J_\alpha$. This implies that $\alpha = \frac{1}{3}$ and, since $c_1 + c_2 + c_3 \geq 1$, $J_0 = \emptyset$ and $|J_{\frac{1}{3}}| \geq 3$. If $J_{\frac{1}{2}} = \emptyset$ then we get (7) with $(I_1, I_2, I_3) = (J_{\frac{1}{3}}, J_{\frac{2}{3}}, J_1)$. If $J_{\frac{1}{2}} \neq \emptyset$, then we get (9) with $(I_2, I_3, I_4, I_6) = (J_{\frac{1}{3}}, J_{\frac{1}{2}}, J_{\frac{2}{3}}, J_1)$.

Suppose $\{i, j\} \subseteq J_\alpha$ and $k \in J_{\frac{1}{2}}$. This implies $\alpha = \frac{1}{4}$, and since $c_1 + c_2 + c_3 \geq 1$, we have $J_0 = \emptyset$, $|J_\alpha| = 2$, and $J_{\frac{1}{2}} \geq 1$. So, we get (8) with $(I_1, I_2, I_3, I_4) = (J_{\frac{1}{4}}, J_{\frac{1}{2}}, J_{\frac{3}{4}}, J_1)$.

Suppose $\{i, j\} \subseteq J_\alpha$ and $k \in J_\beta$. This implies $2\alpha + (1 - \alpha) = 1$, and so $\alpha = 0$. This contradicts $\alpha > 0$.

Finally if $|\{i, j, k\} \cap J_\alpha| \leq 1$, then $c_i + c_j + c_k > 1$, which is a contradiction. \square

Applying Theorem 1, we obtain the following result.

Corollary 13. *Let $S \subseteq \{0, 1\}^n$ be a set with pitch at most 3. Then the CG-rank of every polytope $R \subseteq [0, 1]^n$ with $R \cap \{0, 1\}^n = S$ is at most 8.*

Note that when $\text{tw}(H[\bar{S}]) \leq 2$, none of the inequalities (7), (8), or (9) can appear in the linear description of $\text{conv}(S)$ because for each of them there is a set of indices $I \subseteq [n]$ of size 3 such that the characteristic vector of every proper subset of I is in \bar{S} . This implies that $H[\bar{S}]$ contains a subdivision of K_4 . Hence, we recover the same upperbound of 4 on the CG-rank when $\text{tw}(H[\bar{S}]) \leq 2$ established by Cornuéjols & Lee [8]. On the other hand, the pitch 3 case includes graphs of unbounded treewidth. For example, if we let $S \subseteq \{0, 1\}^n$ be the set of vectors of support at least 3, then S has pitch 3 and $H[\bar{S}]$ contains a subdivision of K_{n+1} .

6. EXTENDED FORMULATIONS VIA SMALL-DEPTH CIRCUITS

Let $S \subseteq \{0, 1\}^n$ be a set with pitch $p = \mathcal{O}(1)$. By Proposition 4, we can efficiently solve linear programs over $P = \text{conv}(S)$, provided that membership in S can be decided efficiently. Here we discuss sizes of extended formulations for such polytopes P . The *extension complexity* $\text{xc}(P)$ of P is defined as the minimum number of facets of a polytope Q for which there exists an affine map π with $\pi(Q) = P$. Below, we review some basic facts concerning extended formulations. The interested reader may consult [5] or [14] for a more in depth treatment.

Since in our case P has at most 2^n vertices, it is easy to see that $\text{xc}(P) \leq 2^n$ holds. However, even for bounded pitch, we can have $\text{xc}(P) = 2^{\Omega(n)}$. This follows from the counting argument of Rothvoß [16]. Since there exist at least $2^{2^{n-1}}$ pitch-1 sets $S \subseteq \{0, 1\}^n$ (for instance, take S to be any superset of $\{x \in \{0, 1\}^n \mid \sum_{i=1}^n x_i = 1 \pmod{2}\}$), some of the polytopes $P = \text{conv}(S)$ have high extension complexity, see also [1, Thm. 1].

Thus, in order to guarantee extended formulations of smaller size, we need to impose further conditions on the set S . In this part, we restrict ourselves to sets $S \subseteq \{0, 1\}^n$ that also can be decided by a small-depth Boolean circuit. Our result is the following.

Theorem 14. *Let $S \subseteq \{0, 1\}^n$ with pitch p such that there exists a depth- D Boolean circuit (with AND and OR gates of fan-in 2, and NOT gates of fan-in 1) that decides S . Then the extension complexity of $P = \text{conv}(S)$ is $\mathcal{O}(n \cdot 2^{pD})$.*

One consequence of Theorem 14 is that $P = \text{conv}(S)$ has a polynomial-size extended formulation whenever S has constant pitch and can be decided by a boolean circuit of logarithmic depth. Furthermore, it shows that finding an *explicit* family of sets $S \subseteq \{0, 1\}^n$ with constant

pitch such that $\text{xc}(\text{conv}(S))$ grows exponentially in n would yield an *explicit* Boolean function admitting no small Boolean circuit of depth $D < \varepsilon n$ for some $\varepsilon > 0$, which appears to be a widely open problem in circuit complexity.

Next, we discuss two tools used in the proof of Theorem 14.

First, we rely on Yannakakis' factorization theorem [19]: If $P = \text{conv}(S) = \{x \mid Ax - b \geq 0\}$ is neither empty nor a point, then $\text{xc}(P)$ equals the *nonnegative rank* of its *slack matrix* M , where the j -th column of M equals the value of $Ax - b$ when replacing x by the j -th element of S , and the nonnegative rank of M is the minimum number of nonnegative rank-1 matrices that sum up to M .

Second, we use a lemma, see Lemma 15 below, which enables us to partition the slack matrix M into submatrices with nonnegative rank $\mathcal{O}(n)$. This trivially implies that the nonnegative rank of the whole matrix M is $\mathcal{O}(nt)$, where t is the number of submatrices in the partition. The lemma implicitly relies on *Karchmer-Wigderson games* [15], a fundamental concept linking communication complexity and circuit complexity which has inspired two recent results on extended formulations [13, 2]. Roughly speaking, our partition of the slack matrix relies on the solution of at most p Karchmer-Wigderson games.

Lemma 15. *Let $S \subseteq \{0, 1\}^n$ such that there exists a depth- D Boolean circuit that decides S . Then there exists a finite set Ω , a relation $K \subseteq \Omega \times \{0, 1\}^n$, and a function $c : \Omega \rightarrow [n]$ such that:*

- (i) $|\Omega| \leq 2^D$,
- (ii) *for every $(s, \bar{s}) \in S \times \bar{S}$ there is exactly one $\omega = \omega(s, \bar{s}) \in \Omega$ with $\omega K s$ and $\omega K \bar{s}$, which we call the witness for (s, \bar{s}) .*
- (iii) *for every $(s, \bar{s}) \in S \times \bar{S}$ with witness $\omega = \omega(s, \bar{s})$, the vectors s and \bar{s} differ at coordinate $c(\omega)$.*

Proof. For each gate of the circuit that is an AND or an OR gate, we label the two input wires of the gate arbitrarily with distinct labels in $\{0, 1\}$. We call a word $\omega \in \{0, 1\}^*$ a *trace* if it encodes a walk in the circuit starting at the output wire and going backwards up the circuit to an input wire. The i th bit of ω defines the label of the input wire of the i th AND or OR gate encountered, that is followed by the walk. When the walk arrives at a NOT gate, it simply continues by following its unique input wire. For each trace ω , let $c(\omega) \in [n]$ denote the index of the input bit reached by ω .

Feeding any input $x \in \{0, 1\}^n$ to the circuit determines a value in $\{0, 1\}$ for each wire (not to be confused with the label of the wire). We say that a trace $\omega \in \{0, 1\}^*$ and an input $x \in \{0, 1\}^n$ to the circuit are *consistent* if at each AND gate or OR gate g that ω encounters, the trace ω follows the *lowest-label* input wire that forces the value of the output of g for input x , provided such an input wire exists. That is, if g is an AND gate (resp. OR gate) with output 0 (resp. 1) for x , we require that ω follows the input wire with a 0-value (resp. 1-value) that has the lowest label. Otherwise, ω may choose the next wire arbitrarily.

Let Ω be the set of traces. Obviously, the length of a trace is at most the depth D of the circuit, and thus $|\Omega| \leq 2^D$. Let $K \subseteq \Omega \times \{0, 1\}^n$ be the set of consistent pairs (ω, x) , so we have $\omega K x$ if and only if ω and x are consistent. We observe that for each $(s, \bar{s}) \in S \times \bar{S}$, there is a unique trace $\omega \in \Omega$ that is consistent with both s and \bar{s} . This is due to the fact that the values of every wire followed by ω are distinct under inputs s and \bar{s} . Hence, at each gate g traversed by ω , there is a unique input wire that ω is allowed to follow. Moreover, we have $s_{c(\omega)} \neq \bar{s}_{c(\omega)}$. Therefore, Ω , K and c satisfy all the requirements of the lemma. \square

Proof of Theorem 14. Let M be a slack matrix of $\text{conv}(S)$ whose rows correspond to linear inequalities valid for S and whose columns correspond to the points in S . Throughout the proof, we write inequalities that are valid for S as $F \equiv \sum_{i \in I_F} c_i^F x_i + \sum_{j \in J_F} c_j^F (1 - x_j) \geq \delta_F$ where $c^F \geq 0$, $\delta_F \geq 0$, and I_F, J_F are a partition of $[n]$, i.e., $I_F \cup J_F = [n]$ and $I_F \cap J_F = \emptyset$. First, observe that the submatrix of M consisting of all trivial rows (that is, those with $\delta_F = 0$) has nonnegative rank at most $2n$. Thus, we may assume that M only contains nontrivial rows.

To each nontrivial inequality F we associate some point $\bar{s}^F \in \{0, 1\}^n$ defined via $\bar{s}_i^F = 0$ if $i \in I_F$ and $\bar{s}_j^F = 1$ if $j \in J_F$. Note that \bar{s}^F is contained in \bar{S} since it violates F .

Let Ω, K, c be as in Lemma 15. For every $\omega_1, \dots, \omega_k \in \Omega$ with $k \leq p$ consider the following submatrix $M_{(\omega_1, \dots, \omega_k)}$ of M . For easier notation, for a point $a \in \{0, 1\}^n$ and any $i \in [k]$ let $a^{(i)}$ denote the vector that arises from a by flipping the coordinates $c(\omega_1), \dots, c(\omega_i)$.

- A column associated to a point $s \in S$ belongs to $M_{(\omega_1, \dots, \omega_k)}$ if and only if:
 - $\omega_i K s$ for every $i = 1, \dots, k$
- A row associated to an inequality F with $\bar{s} := \bar{s}^F$ belongs to $M_{(\omega_1, \dots, \omega_k)}$ if and only if:
 - $\omega_i K \bar{s}^{(i)}$ for every $i = 1, \dots, k$
 - $\bar{s}^{(i)}$ violates inequality F for every $i = 1, \dots, k - 1$
 - $\bar{s}^{(k)}$ satisfies inequality F

From Lemma 15 (ii) one obtains that every entry of M is covered by at most one of the above matrices. On the other hand, we claim that every entry of M is covered by one such matrix $M_{(\omega_1, \dots, \omega_k)}$ with $k \leq p$. To see this, consider any entry (s, F) with $\bar{s} := \bar{s}^F$ and let $\omega_1 \in \Omega$ be the unique witness for (s, \bar{s}) . By Lemma 15 (iii) the vectors s and \bar{s} differ at coordinate $c(\omega_1)$. If the vector $\bar{s}^{(1)}$ obtained from \bar{s} by flipping coordinate $c(\omega_1)$ is valid for F , this entry is covered by $M_{(\omega_1)}$. If not, observe that $\bar{s}^{(1)} \in \bar{S}$ and let ω_2 be the unique witness for $(s, \bar{s}^{(1)})$. If the vector $\bar{s}^{(2)}$ obtained from $\bar{s}^{(1)}$ by flipping coordinate $c(\omega_2)$ is valid for F , this entry is covered by $M_{(\omega_1, \omega_2)}$, and so on. Observe that after every iteration of the described process, we have that s and $\bar{s}^{(i)}$ coincide at coordinates $c(\omega_1), \dots, c(\omega_i)$ and hence these coordinates need to be pairwise distinct. Thus, since the pitch of S is p , we know that $\bar{s}^{(i)}$ is valid for F whenever $i \geq p$.

This shows that every entry of M is covered by exactly one of the above submatrices. By Lemma 15 (i) we need at most $1 + 2^D + \binom{2^D}{2} + \dots + \binom{2^D}{p} \leq 2^{pD+1} - 1 = \mathcal{O}(2^{pD})$ submatrices (including that corresponding to the trivial inequalities) and hence it suffices to show that the nonnegative rank of any submatrix is $\mathcal{O}(n)$. To see this, fix $\omega_1, \dots, \omega_k \in \Omega$ and consider the submatrix $M' := M_{(\omega_1, \dots, \omega_k)}$. Let F and s correspond to a row and column of M' , respectively. Denoting by $T := \{c(\omega_1), \dots, c(\omega_k)\}$, the value of M' at the entry (F, s) is equal to

$$\begin{aligned} & \sum_{i \in I_F} c_i^F s_i + \sum_{j \in J_F} c_j^F (1 - s_j) - \delta_F \\ &= \sum_{i \in I_F \setminus T} c_i^F s_i + \sum_{j \in J_F \setminus T} c_j^F (1 - s_j) + \sum_{i \in I_F \cap T} c_i^F s_i + \sum_{j \in J_F \cap T} c_j^F (1 - s_j) - \delta_F \\ &= \sum_{i \in I_F \setminus T} c_i^F s_i + \sum_{j \in J_F \setminus T} c_j^F (1 - s_j) + \underbrace{\sum_{i \in T} c_i^F}_{\geq 0} - \delta_F, \end{aligned}$$

where the second equality follows from the fact that s and \bar{s} differ in every coordinate in T , so that $s_i = 1$ for $i \in I_F$ and $s_j = 1$ for $j \in J_F$, and the nonnegativity of the last term follows from the fact that flipping all the coordinates of \bar{s} that are in T results in a vector that is valid for F . This shows that M' can be written as the sum of at most $n - |T| + n - |T| + 1 \leq 2n$ nonnegative rank-1 matrices. \square

7. ACKNOWLEDGMENTS

We thank Pierre Aboulker, Yuri Faenza and Robert Weismantel for discussions at an early stage of this research. We also thank Gennadiy Averkov, Michele Conforti and Volker Kaibel for their feedback. Yohann Benchetrit, Samuel Fiorini and Tony Huynh were supported by ERC Consolidator Grant 615640-ForEFront.

REFERENCES

- [1] Gennadiy Averkov, Voker Kaibel, and Stefan Weltge. Maximum semidefinite and linear extension complexity of families of polytopes, 2016.

- [2] Abbas Bazzi, Samuel Fiorini, Sangxia Huang, and Ola Svensson. Small extended formulation for knapsack cover inequalities from monotone circuits. In *Proc. SODA 2017*, 2017.
- [3] Daniel Bienstock and Mark Zuckerberg. Subset algebra lift operators for 0-1 integer programming. *SIAM J. Optim.*, 15(1):63–95, 2004.
- [4] Vašek Chvátal, William Cook, and Mark Hartmann. On cutting-plane proofs in combinatorial optimization. *Linear algebra and its applications*, 114:455–499, 1989.
- [5] M. Conforti, G. Cornuéjols, and G. Zambelli. Extended formulations in combinatorial optimization. *4OR*, 8:1–48, 2010.
- [6] Michele Conforti, Gérard Cornuéjols, and Giacomo Zambelli. *Integer Programming (Graduate Texts in Mathematics)*. Springer, 2014.
- [7] Michele Conforti, Alberto Del Pia, Marco Di Summa, Yuri Faenza, and Roland Grappe. Reverse Chvátal–Gomory rank. *SIAM Journal on Discrete Mathematics*, 29(1):166–181, 2015.
- [8] Gérard Cornuéjols and Dabeen Lee. On some polytopes contained in the 0,1 hypercube that have a small Chvátal rank. In *International Conference on Integer Programming and Combinatorial Optimization*, pages 300–311. Springer, 2016.
- [9] Gérard Cornuéjols and Yanjun Li. Deciding emptiness of the Gomory–Chvátal closure is NP-complete, even for a rational polyhedron containing no integer point. In *International Conference on Integer Programming and Combinatorial Optimization*, pages 387–397. Springer, 2016.
- [10] Daniel Dadush, Santanu S. Dey, and Juan Pablo Vielma. On the Chvátal–Gomory closure of a compact convex set. *Mathematical Programming*, 145(1-2):327–348, 2014.
- [11] Reinhard Diestel. *Graph theory*, volume 173 of *Graduate Texts in Mathematics*. Springer, Heidelberg, fourth edition, 2010.
- [12] Friedrich Eisenbrand and Andreas S. Schulz. Bounds on the Chvátal rank of polytopes in the 0/1-cube. *Combinatorica*, 23(2):245–261, 2003.
- [13] Mika Göös, Rahul Jain, and Thomas Watson. Extension complexity of independent set polytopes. In *Proc. FOCS 2016*, 2016.
- [14] V. Kaibel. Extended formulations in combinatorial optimization. *Optima*, 85:2–7, 2011.
- [15] M. Karchmer and A. Wigderson. Monotone circuits for connectivity require super-logarithmic depth. *SIAM J. Discrete Math.*, 3:255–265, 1990.
- [16] T. Rothvoß. Some 0/1 polytopes need exponential size extended formulations. *Math. Programming A*, 142:255–268, 2013.
- [17] Thomas Rothvoß and Laura Sanità. 0/1 polytopes with quadratic Chvátal rank. In *Integer programming and combinatorial optimization*, volume 7801 of *Lecture Notes in Comput. Sci.*, pages 349–361. Springer, Heidelberg, 2013.
- [18] M. Singh and K. Talwar. Improving integrality gaps via Chvátal–Gomory. In *Approximation, randomization, and combinatorial optimization*, volume 6302 of *Lecture Notes in Comput. Sci.*, pages 366–379. Springer, Berlin, 2010.
- [19] Mihalis Yannakakis. Expressing combinatorial optimization problems by linear programs. *J. Comput. System Sci.*, 43(3):441–466, 1991.